# Chapter 3
# The Deep Lexical Semantics of Emotions

**Jerry R. Hobbs and Andrew S. Gordon**

## 3.1 Introduction

We understand discourse so well because we know so much. If we are to have natural language understanding systems that are able to deal with texts with emotional content, we must encode knowledge of human emotions for use in the systems. In particular, we must equip the system with a formal version of people's implicit theory of how emotions mediate between what they experience and what they do, and rules that link the theory with words and phrases in the emotional lexicon.

The effort we describe here is part of a larger project in knowledge-based natural language understanding to construct a collection of abstract and concrete core formal theories of fundamental phenomena, geared to language, and to define or at least characterize the most common words in English in terms of these theories [8]. One collection of theories we have put a considerable amount of work into is a commonsense theory of human cognition, or how people think they think [9]. A formal theory of emotions is an important piece of this. In this paper we describe this theory and our efforts to define a number of the most common words about emotions in terms of this and other theories.

Vocabulary related to emotions has been studied extensively within the field of linguistics, with particular attention to cross-cultural differences [1, 6, 18]. Within computational linguistics, there has been recent interest in creating large-scale text corpora where expressions of emotion and other private states are annotated [17].

In Section 3.2 we describe Core WordNet and our categorization of it to determine the most frequent words about cognition and emotion. In Section 3.3 we describe an effort to flesh out the emotional lexicon by searching a large corpus for emotional terms, so we can have some assurance of high coverage in both the core theory and the lexical items linked to it. In Section 3.4 we sketch the principal facets of some of the core theories. In Section 3.5 we describe the theory of Emotion with several examples of words characterized in terms of the theories.

J.R. Hobbs (✉)
Information Sciences Institute, University of Southern California, Marina del Rey, CA 90292, USA
e-mail: hobbs@isi.edu

K. Ahmad (ed.), *Affective Computing and Sentiment Analysis*, Text, Speech and Language Technology 45, DOI 10.1007/978-94-007-1757-2_3,

## 3.2 Identifying the Core Emotion Words

WordNet [14, 15] contains tens of thousands of synsets referring to highly specific animals, plants, chemical compounds, French mathematicians, and so on. Most of these are rarely relevant to any particular natural language understanding application. To focus on the more central words in English, the Princeton WordNet group has compiled a CoreWordNet, consisting of 4,979 synsets that express frequent and salient concepts. These were selected as follows: First, a list with the most frequent strings from the British National Corpus was automatically compiled and all WordNet synsets for these strings were pulled out. Second, two raters determined which of the senses of these strings expressed "salient" concepts [2]. CoreWordNet is downloadable from http://wordnet.cs.princeton.edu/downloads.html.

Only nouns, verbs and adjectives were identified in that effort, but subsequently 322 adverbs were added to the list.

We classified these word senses manually into sixteen broad categories, including such classes as Composite Entities, Scales, Events, Space, Time, Communication, Microsocial (e.g., personal relationships), Macrosocial (e.g., government), Artifacts, and Economics. A very important class was Cognition, or concepts involving mental and emotional states. This included such words as *imagination, horror, rely, remind, matter, estimate*, and *idea*. Altogether 778 words senses were put into this class.

These were further divided into thirty classes based on commonsense theories of cognition we had identified from an examination of several hundred human strategies [4] and had constructed formal theories of in a defeasible, first-order predicate calculus [9]. Among the thirty are theories of Knowledge Management, Memory, Goals and Plans, Envisionment (or "thinking about"), Decisions, Threat Detection, Explanations, and Emotions. 140 of the 778 cognitive word senses concern emotions, and are the focus of this paper. Some random examples of the emotion word senses are as follows (many of these are ambiguous, but it is the emotional sense that concerns us): *heart, concern, relief, anger, mood, joy, fit, embarrassment, morale, apathy, pride, disgust, want, feel, suffer, cry, upset, provoke, terrify, fascinate, glad, exciting, happy, sympathetic, passionate*, and *calmly*.

## 3.3 Filling Out the Lexicon of Emotion

With the aim of providing automated tools for annotating expressions of emotion in English text, we developed a catalogue of English words and phrases that refer to emotional states and emotion-related mental events, as part of a larger effort to recognize all English expressions related to commonsense psychology [5].

Our strategy consisted of three steps. First, we convened a group brainstorming meeting with researchers, graduate students, and administrative staff within our research lab. Participants were asked to creatively and competitively produce words and phrases that were related to emotional states, the expression of emotions, and commonsense mental processes involving emotions. The purpose of this meeting

was to produce an initial list that could serve as the starting point for an exhaustive linguistic search. Second, a team of graduate students in linguistics and computational linguistics were tasked to elaborate this list by consulting a variety of thesauri, phrase dictionaries, and electronic linguistic resources. WordNet was particularly useful during this step; the list was expanded to include all hyponyms of *emotion*-1, troponyms of *provoke*-1, and troponyms of *feel*-1. Morphological derivatives of each word in the expanded list were also included, e.g., the verb *resent* relates both to its present participle (*resenting*), but also to the adjective *resentful* and its derivatives (*resentfully* and *resentment*). Third, the resulting list (several hundred emotion terms) was then organized into semantic classes by clustering terms with similar meaning. During this step, we relied heavily on the emotion categories proposed by Ortony et al. [16], expanded by Clark Elliott [3] to include 24 distinct emotion types. The final taxonomy added a superordinate emotion class, a class for the lack of emotion, and seven classes of terms related to emotion-related mental processes, resulting in a final list of 33 taxonomic distinctions.

In conducting this analysis, we were particularly struck by two characteristics of emotion vocabulary that distinguishes it from other terminology related to commonsense psychology, e.g. beliefs, goals and plans. First is the sheer quantity of single words that reference emotion states in the English language, in no small part due to the borrowing power of English; there are literally hundreds of words available to English-speakers to describe how they are feeling. Second is the low level of polysemy within this set; most emotion terms have only a single word sense. The list below provides several examples of each of the 33 emotion categories, with the adjectival form favored over other derivatives.

1. emotion (*affect, emotion, feeling, have feelings of*)
2. joy emotion (*blithe, cheery, comfortable, ecstatic, elated, enjoyment, happy, be in high spirits, be in Nirvana, be on cloud nine*)
3. distress emotion (*agony, bereavement, brokenhearted, cheerless, depression, despondent, sad, tearful, unhappy, be low spirited, have a sinking feeling*)
4. happy-for emotion (*glad for, pleased for, congratulatory*)
5. sorry-for emotion (*commiserative, compassionate, condolence*)
6. resentment emotion (*covetous, envious, jealous, sulky, vengeful*)
7. gloating emotion (*schadenfreude, mawkish*)
8. hope emotion (*encouragement, hopeful, optimistic, sanguine*)
9. fear emotion (*anxious, apprehensive, bode, consternation, despair, fearful, terror, timid, trepidation, uneasy, worried, have cold feet, gives one the creeps*)
10. satisfaction emotion (*consolation, delightful, gratification, pleasure, ravishment, satisfaction, solace, have a silver lining*)
11. fears confirmed emotion (*fears have come true, fears realized*)
12. relief emotion (*alleviation, assuagement, relief*)
13. disappointment emotion (*defeat, disappointment, frustration*)
14. pride emotion (*conceited, egotistic, proud, prideful, vain*)
15. self-reproach emotion (*chagrin, discomfit, embarrassment, humble, humility, meek, repentance, self-conscious, self-depreciation, shame*)

16. appreciation emotion (*appreciative, thankful*)
17. reproach emotion (*disapproval, reproachful*)
18. gratitude emotion (*grateful*)
19. anger emotion (*aggravation, angry, annoyance, belligerent, furious, pique, rage*)
20. gratification emotion (*gratifying*)
21. remorse emotion (*guilt, regretful, remorseful, rueful*)
22. liking emotion (*fancy, fascination, fondness, partiality, penchant, predilection, have a taste for, have a weakness for*)
23. disliking emotion (*abhorrent, abomination, detestable, disinclination, dislikable, execration, loathsome, repugnant, repulsive, revulsion*)
24. love emotion (*adoration, agape, amorous, devotion, enamor, infatuation, lovable*)
25. hate emotion (*animosity, bitterness, despise, hateful, malefic, malevolent, malicious, spite, venomous, have bad blood*)
26. emotional state (*mood, way one feels, how one is feeling*)
27. emotional state explanation (*reason for feeling, why one feels, cause of the emotion*)
28. emotional state change (*a shift in mood*)
29. appraisal (*assess one's emotions, figure out how one feels about*)
30. coping strategy (*way of dealing with, coping technique*)
31. coping (*dealing with the feeling, coming to terms with*)
32. emotional tendency (*emotional, moodiness, passionate, sentimentality*)
33. no emotion (*aloof, ambivalent, austere, calm, cold-hearted, emotionless, heartless, impassive, indifferent, phlegmatic*)

## 3.4 Some Core Theories

We use first-order logic for encoding axioms in our commonsense theories, in the syntax of Common Logic [13]. Since human cognition concerns itself with actual and possible events and states, which we refer to as eventualities, we reify these and treat them in the logic as ordinary individuals. Similarly, we treat sets as ordinary individuals and axiomatize naive set theory. Most axioms are only normally true, and we thus have an approach to defeasibility – proofs can be defeated by better proofs. Our approach to defeasibility is based on weighted abduction [11] and is similar to McCarthy's circumscription [12], but the content of the theories should survive a translation to any other adequate framework for defeasibility.

The theories of cognition rest on sixteen background theories. Included among these is a theory of scales that provides means of talking about partial orderings, the figure-ground relation of placing some external thing *at* a point on a scale, and qualitative regions identifying the *high* and *low* regions of a scale. The latter are linked to the theory of functionality mentioned below; often when we call something tall, we mean tall enough for some purpose. They also need to be linked to an as-yet undeveloped commonsense theory of distributions.

In addition, we have theories of change of state, causality, and time. The theory of causality tries to provide a defeasible notion of *cause* that can be used in lexical semantics [7]. The theory of time explicates such predicates as *before*, *atTime* relating an event to a time, and a *meets* relation between intervals [10].

For this paper the most relevant cognitive theories are Knowledge Management, Goals and Planning, and Envisionment. In the theory of Knowledge Management, we characterize belief and graded belief and their relation to perception, inference, and action. Briefly, perceiving is believing, we can defeasibly do logic inside belief contexts, and our beliefs influence our actions. We also axiomatize change of belief, mutual belief, assuming, varieties of inference, justification, knowledge domains, expertise, and other similar concepts in this theory.

The theory of Goals and Planning posits agents that have a top-level goal "to thrive", have various beliefs about what will cause them to thrive and other causal knowledge, and continually plan and replan to achieve this top-level goal. Planning uses axioms about what eventualities cause or enable what other eventualities to generate subgoals of goals, and subgoals of the subgoals, until arriving at executable actions. Shared goals and plans are defined in terms of mutual knowledge and of sets of agents having goals where the shared plans bottom out in actions by individual members. We define notions of eventualities being good for or bad for an agent or group of agents relative to their goals. The function and roles of artifacts and organizations are characterized in terms of agents' goals, where the structure of the artifact or organization reflects the structure of the plan to achieve the goals. We also explicate here the notions of attempting to achieve a goal and actually achieving it. A *threat* is an eventuality that may cause one's goals not to be achieved.

The theory of Envisionment is an attempt to begin to capture what it is to think about something, particularly, in a causal manner. To envision is to entertain in one's focus of attention a sequence of causally linked sets of eventualities. For example, the Common Logic expression `(envisionFromTo a s1 s2)` says that an agent `a` envisions a sequence of causally connected situations starting with `s1` and ending with `s2`. Explanation, prediction, and planning are varieties of envisionment.

## 3.5 The Theory and Lexical Semantics of Emotion

Our theory of Emotions attempts to characterize twenty-six basic emotions in terms of the abstract situations that cause them and the abstract classes of behavior they trigger. That is, emotions are viewed primarily as mediating between perception and action. Our treatment is based in part, but only in part, on that of Ortony et al. [16]. We attempt, in addition, to axiomatize the notion of the *intensity* of emotion, and give a somewhat more central role to the "raw emotions", as described below.

Natural language is very rich in emotional terminology, and our formal theory of emotion tracks language very closely. Thus, in explicating the concepts of the theory, we are also providing the deep lexical semantics of English emotional terms. Of course, the converse is not also true; there are many more English emotional

terms than would be basic predicates in an underlying theory of emotion; these others we characterize in terms of the basic predicates.

*Happiness* is normally caused by the belief that one's goals are being satisfied. This of course is not always the explanation of one's happiness. Imagining you will win the lottery can cheer you up, sometimes you feel happy for no identifiable reason at all, and sometimes you are unhappy even though everything is going well. This is an illustration of why virtually all the rules in the cognitive theories are defeasible.

To give a flavor of the rules in the theories, we include the fairly complex one characterizing one of the sources of happiness.

```
(forall (a g e1 e2 e3 t1 t2)
  (if (and (goal' e1 g a)
           (atTime e1 t1)
           (atTime' e2 g t2)
           (believe' e3 a e2)
           (atTime e3 t1)
           (intMeets t1 t2) <etc>)
      (exists (e4)
         (and (happy' e4 a)
              (atTime e4 t1)
              (cause e3 e4)))))
```

That is, if during time interval `t1` agent `a` has the goal `g` and believes that it will be satisfied during interval `t2`, where `t2` begins when `t1` ends, then this belief will cause `a` to be happy during interval `t1`. More succinctly, anticipating success makes us happy. The `<etc>` is an abbreviation indicating defeasibility.

An inference one can draw from one's success in satisfying one's goals is that the rules or beliefs that generate one's behavior are functional. They are the right rules. Therefore, there are two conclusions with respect to one's actions. Since the rules are correct, there will be a reluctance to change one's beliefs, at least in the relevant knowledge domains. The current beliefs are doing a good job. And one will be inclined to act on one's current beliefs. One will exhibit a greater level of activity.

Sadness is given a corresponding characterization. It is normally caused by the belief that one's goals are not being satisfied. It tends to suppress the urge to action, since one would be acting on beliefs that have shown themselves to be dysfunctional. Moreover, sadness opens one to a change in beliefs.

We have axiomatized Ortony et al.'s [16] cognitive elaborations on basic emotions. Happiness and sorrow for someone else, resentment, and gloating are defined in terms of eventualities being good for or bad for in-groups and out-groups, where in-groups are defined in terms of shared goals. Anticipation is defined in terms of envisionment; satisfaction, "fears confirmed", disappointment, and relief are defined in terms of anticipated eventualites that are good for or bad for the agent being realized or frustrated. Pride, self-reproach, appreciation, reproach, gratification, remorse, gratitude and a certain kind of anger are defined in terms of eventualities that are good for or bad for one's self or others being merely attempted or succeeding.

Although we do not "define" emotional intensity, we do constrain its interpretations with axioms that say in some special circumstances what sort of emotions will normally be more intense than others, *ceteris paribus*. For example, normally the more salient the stimulus, the more intense the emotion, and the more intense the emotion the more extreme the response. *Intense* then labels the functionally and distributionally high region of that scale.

Our treatment of the three "raw" emotions, anger, fear, and disgust, depends on the notions of eliminating or avoiding threats. One eliminates a threat by causing a change of state (or location) in it. One avoids a threat by causing a change of state (or location) in one's self. In either case, the effect is a reduction of the threat. Anger and fear are both caused by threats. In anger, our response to it is normally to try to eliminate the threat. In fear, our response is normally to try to avoid the threat.

Fear and anger are responses to external threats. Disgust is a response to a threat that is interior, and it triggers an effort to eject the threat. "Interior" may be interpreted literally with respect to the body – most of the ways of talking about disgust involve distaste or nausea. Or we may interpret it metaphorically as referring to an in-group.

All of this is of course quite naive if viewed as a *real* theory of emotions. But we believe it is reasonable as a commonsense theory, and will allow natural language systems to make sense of most occurrences of emotion terms in English discourse.

Having explicated the basic emotions formally, we are now able to write axioms characterizing the meanings of the less central emotional terminology of English. For example, to "terrify" someone is to cause one to feel intense fear. The various emotional word senses of "calm" in WordNet can be characterized in terms of feeling or causing low emotional intensity.

There are five noun senses of *pride* in WordNet. *pride-N2* includes the Ortony et al.'s [16] sense we characterized above as what one feels on an attempt to do something good, but also includes the feeling on success and the feeling about another person's attempt or success. *pride-N1* is a version of *pride-N2*, generalized over time. *pride-N3* refers to the causal power of *pride-N1* in one's actions. *pride-N5* is *pride-N1* carried to excess. (The fourth sense is a group of lions.) The single verb sense of *pride* means to feel or express *pride-N1*.

## 3.6 Summary

Natural language understanding requires a large knowledge base of commonsense knowledge that explicates concepts in coherent theories and links lexical items with these theories. In order to achieve high accuracy, high complexity results, this effort must be manual (as indeed dictionaries are constructed manually). Early efforts will have the most impact if done for the most central concepts and the most common word senses.

In this paper we have outlined our work in constructing background theories and theories of general cognition, and we have described in more detail the structure

of the theory of Emotion, indicating how it can be used to explicate the emotional vocabulary of English.

## References


1. Athanasiadou, Angeliki and Elzbieta Tabakowska. eds. 1998. *Speaking of emotions: Conceptualisation and expression*. Berlin, NY: Mouton de Gruyter.
2. Boyd-Graber, Jordan, Christiane Fellbaum, Dan Osherson, and Robert Schapire. 2006. Adding dense, weighted, connections to WordNet. Proceedings, Third Global WordNet Meeting, Jeju Island, Korea, January 2006.
3. Elliott, Clark. 1992. The affective reasoner: A process model of emotions in a multi-agent system. PhD diss., Northwestern University. The Institute for the Learning Sciences, Technical Report No. 32.
4. Gordon, Andrew S. 2004. *Strategy representation: An analysis of planning knowledge*. Mahwah, NJ: Lawrence Erlbaum Associates.
5. Gordon, Andrew S., Abe Kazemzadeh, Anish Nair, and Milena Petrova. 2003. Recognizing expressions of commonsense psychology in English text. Proceedings, 41st annual meeting of the Association for Computational Linguistics (ACL-2003), Sapporo, Japan, July 2003.
6. Harkins, Jean, and Anna Wierzbicka, eds. 2001. *Emotions in crosslinguistic Perspective*. Berlin, NY: Mouton de Gruyter.
7. Hobbs, Jerry R. 2005. Toward a useful notion of causality for lexical semantics. *Journal of Semantics* 22:181–209.
8. Hobbs, Jerry R. 2008. Deep lexical semantics. Proceedings, 9th international conference on Intelligent Text Processing and Computational Linguistics, Haifa, Israel, February 2008.
9. Hobbs, Jerry R., and Andrew S. Gordon. 2005. Encoding knowledge of commonsense psychology. Proceedings, 7th international symposium on Logical Formalizations of Commonsense Reasoning, Corfu, Greece, pp. 107–114, May 2005.
10. Hobbs, Jerry R. and Feng Pan. 2004. An ontology of time for the Semantic Web. *ACM Transactions on Asian Language Information Processing* 3:66–85.
11. Hobbs, Jerry R., Mark Stickel, Douglas Appelt, and Paul Martin. 1993. Interpretation as abduction. *Artificial Intelligence* 63:69–142.
12. McCarthy, John. 1980. Circumscription: A form of nonmonotonic reasoning. *Artificial Intelligence* 13:27–39.
13. Menzel, Chris, et al. 2008. Common logic standard. http://cl.tamu.edu/.
14. Miller, George. 1995. WordNet: A lexical database for English. *Communications of the ACM* 38:39–41.
15. Miller, George, Christiane Fellbaum, and Randee Tengi. 2006. WordNet: A lexical database for the English language. http://wordnet.princeton.edu/.
16. Ortony, Andrew, Gerald L. Clore, and Allan Collins. 1988. *The cognitive structure of emotions*. New York: Cambridge University Press.
17. Wiebe, Janyce, Theresa Wilson, and Claire Cardie. 2005. Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation* 39:165–210.
18. Wierzbicka, Anna. 1999. *Emotions across languages and cultures: Diversity and universals*. Cambridge: Cambridge University Press.